

RESEARCH ARTICLE

Simple Sentiment Analysis Ansatz for Sentiment Classification in Quantum Natural Language Processing

FARISKA ZAKHRALATIVA RUSKANDA^{1,2}, (Member, IEEE),
MUHAMMAD RIFAT ABIWARDANI¹, INFALL SYAFALNI^{1,3,4}, (Member, IEEE),
HARASHITA TATIMMA LARASATI^{1,5}, (Graduate Student Member, IEEE),
AND RAHMAT MULYAWAN^{1,3,6}, (Member, IEEE)

¹School of Electrical Engineering and Informatics, Bandung Institute of Technology, Bandung 40132, Indonesia

²Artificial Intelligence Center, Bandung Institute of Technology, Bandung 40132, Indonesia

³University Center of Excellence on Microelectronics, Bandung Institute of Technology, Bandung 40132, Indonesia

⁴Interuniversity Microelectronics Centre (IMEC), 3001 Leuven, Belgium

⁵School of Computer Science and Engineering, Pusan National University, Busan 46241, South Korea

⁶Research Collaboration Center for Quantum Technology 2.0, Bandung Institute of Technology, Bandung 40132, Indonesia

Corresponding author: Infall Syafalni (infall@ieee.org)

This work was supported by the School of Electrical Engineering and Informatics, Bandung Institute of Technology.

ABSTRACT Sentiment classification is a valuable application of natural language processing that has seen wide usage in optimizing business processes. This paper explores a novel implementation of sentiment analysis using the Variational Quantum Algorithms (VQA) framework. As ansatz choice determines model performance in VQA, this paper proposes an alternative ansatz for the sentiment classification task in quantum representation. Specifically, it builds upon previous work in quantum sentiment classification by proposing an alternative ansatz to the Instantaneous Quantum Polytime ansatz, entitled Simple Sentiment Analysis (SimpleSA) ansatz. A key feature of the SimpleSA ansatz is the decision to neglect noun parameterization. The proposed SimpleSA has less complexity than the other ansätze in terms of the number of parameters and the number of gates. Moreover, experimental results show that the SimpleSA ansatz with H-CNOT-Rz-H compound block construction outperforms the Instantaneous Quantum Polytime (IQP) ansatz at 85.00% accuracy. Furthermore, SimpleSA optimization converges 20.89% faster than Instantaneous Quantum Polytime (IQP) for the Simultaneous Perturbation Stochastic Approximation (SPSA) method with 130 iterations. The proposed work is useful for applications of quantum computers for sentiment analyses and classifications.

INDEX TERMS Quantum natural language processing, quantum machine learning, sentiment classification, ansatz, quantum circuit.

I. INTRODUCTION

Recent breakthroughs in quantum computing technology have expanded the scope of multiple disciplines beyond what was previously possible. Among the fields impacted is natural language processing, or NLP. Compared to other branches of machine learning, NLP has seen significant developments with the advent of quantum machine learning (QML), giving

The associate editor coordinating the review of this manuscript and approving it for publication was Agostino Forestiero^{id}.

rise to the field of quantum natural language processing, or QNLP.

The foundations for QNLP were first established in [1]. The paper describes an equivalence between pregroup grammars and diagrams for quantum processes, formulated in a representation dubbed “DisCoCat”. This was understood in [1] as a distributional and compositional model of meaning. The model preserves the syntax of a sentence throughout diagram manipulation and, as such, allows the grammatical properties of sentences to be reflected when calculating a

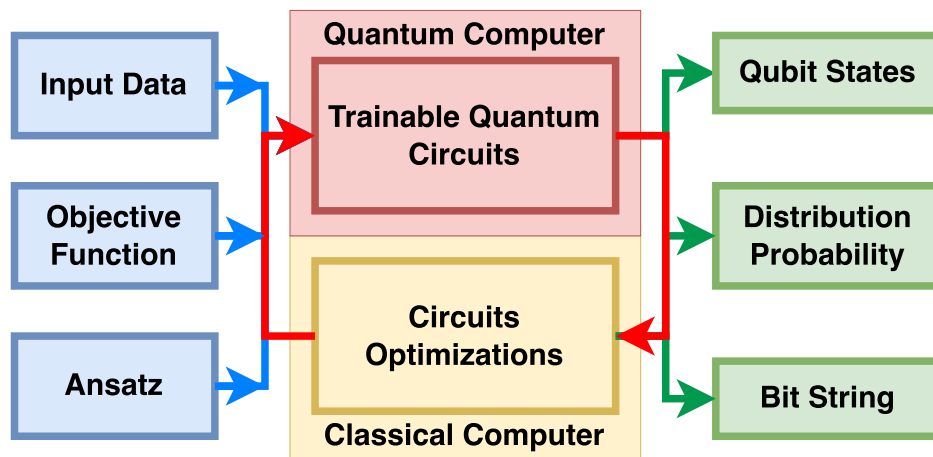


FIGURE 1. Diagram of VQA scheme.

sentence's meaning. Such a model was unorthodox around the late 2000s when most language models employed a bag-of-words approach. Unfortunately, around the time the paper was published, there were no quantum computers capable of simulating the proposed model, which would only arrive almost a decade later.

Since 2018, there have been significant developments in the field of quantum computing. Now, quantum computer technology has entered the "NISQ" era. NISQ stands for Noisy Intermediate-Scale Quantum, where quantum hardware is in the fifty to several hundred qubits range and circuit operations are noisy [2]. Indeed, this class of quantum hardware has been used to implement QML models. In order to mitigate the drawbacks of NISQ-era computers, a hybrid quantum-classical approach to quantum machine learning was introduced, known as Variational Quantum Algorithms, or VQA [3]. Essentially, this framework uses quantum computers to evaluate parameterized quantum circuits (PQCs), then approximates the necessary changes in circuit parameter values using a classical computation method until the model converges. Fig. 1 depicts the general VQA framework, adapted from [3].

As a product of quantum computational development within the past four years, research in the field of QNLP has advanced considerably. The feasibility of implementing NLP using quantum computers was analyzed in [4], which compiles the first full-stack pipeline for DisCoCat QNLP. The paper specifies the stages of data transformation as sentences are converted into string diagrams and quantum circuits. Furthermore, the work improves the string diagram simplification method offered in [5], which is used as groundwork for further QNLP implementation. Then, in [6], a QNLP pipeline based on the principles of VQA was implemented, specifying the data transformation and problem modeling techniques based on the calculus rules described in [7]. Eventually, the open-source Python library, Lambeq [6], was published for accessible DisCoCat QNLP use. The pipeline offered in the work details the stages of transforming

sentences as an input into the quantum machine learning model's prediction as its output. It utilizes the DisCoCat model proposed in [1] and builds trainable quantum circuits based on the string diagram representation, using a class of algorithms known as ansätze.

An ansatz, in the context of quantum computing, is a subroutine or algorithm that dictates how a parameterized quantum circuit is generated from a problem model [8]. In the context of DisCoCat QNLP, the ansatz determines how specific word boxes of string diagrams are transcribed to specific arrangements of quantum circuit gates [6].

The general QNLP pipeline as proposed in [6] is depicted in Fig. 2. In addition to seminal papers that establish the fundamental theories, other researchers have applied QNLP theory to real-life use cases within natural language processing. Among them is [9], [10], and [11], which uses the QNLP framework previously described in order to perform sentiment analysis.

Sentiment analysis is a task within the field of NLP that aims to analyze the opinions or sentiments of a speaker through the use of sentences in natural language [12]. In the industry, sentiment analysis is regularly done to determine public opinion on a product or service offered by a company [13]. In this work, the scope of sentiment analysis is limited to sentence-level binary sentiment classification. While multi-class sentiment classification is possible with DisCoCat QNLP as shown in [14], this research opts for a simplified model in order to streamline the ansatz design process. That is, binary sentiment classification allows the output register to be represented by a single qubit; positive or negative sentiment is assigned to the computational basis states $|0\rangle$ and $|1\rangle$. Such is also done in [9], [10], and [11].

In [9], sentiment classification is performed by measuring a sentence's PQC bit string output and interpreting it as the model's prediction. The paper integrates the DisCoCat model to represent sentences in string diagrams and uses IQP ansatz to transcribe the diagrams into PQCs. The PQCs are optimized classically using SPSA. This is done to speed up

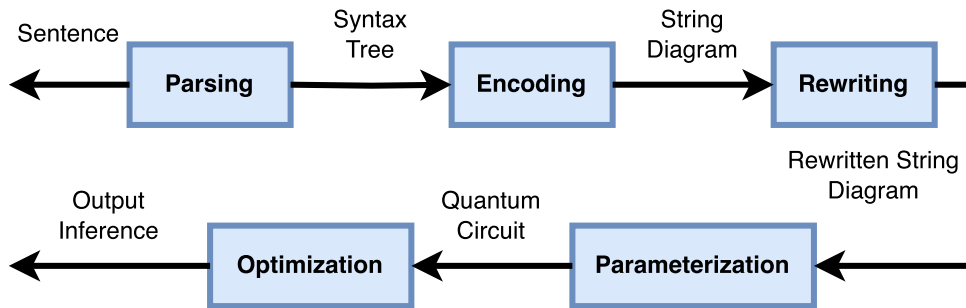


FIGURE 2. QNL pipeline diagram [6].

computation. Even so, the optimization process still took an extensive amount of time. On a simplified artificial dataset, training and classification required an order of 2.5 hours to complete. Additionally, the performance of the model still needs to be improved as the model achieved at most 81.67% on the test set. Therefore, this research aims to find an alternative hybrid quantum-classical machine learning method for sentiment analysis with better performance. In particular, this research is intended as a direct comparison of performance to the model offered in [9]. Specifically, this research implements an alternative ansatz that is better tailored for sentiment classification.

This study aims to address the limitations of the IQP ansatz for sentiment analysis by exploring an alternative ansatz, referred to as the SimpleSA ansatz. Unlike the IQP ansatz, SimpleSA does not parameterize noun types into rotation gates. The study seeks to leverage this alternative ansatz within the VQA framework to enhance the performance of sentiment analysis. Utilizing a hybrid quantum-classical machine learning approach, it aims to achieve better accuracy and efficiency compared to the existing model.

The main contributions of this paper are the following:

- 1) We explore a novel implementation of sentiment analysis for use in quantum natural language processing (QNL) using the Variational Quantum Algorithms (VQA) framework.
- 2) We propose a novel ansatz called Simple Sentiment Analysis (SimpleSA).
- 3) We elaborate the detailed pipeline of Simple Sentiment Analysis (SimpleSA) which consists of sentence-to-diagram conversion, diagram-to-circuit conversion, optimization, and metric measurement.
- 4) We calculate the derivation of the wave function and the complexity of the proposed SimpleSA ansatz in term of number of parameters and gates. Moreover, we also analyze the accuracy of the proposed SimpleSA ansatz.
- 5) Finally, our proposed method (SimpleSA ansatz) with H-CNOT-Rz-H compound block construction outperforms IQP ansatz at 85.00% accuracy. Furthermore, SimpleSA optimization converges 20.89% faster than IQP for SPSA with 130 iterations.

The remainder of this paper is organized as follows: First, we provide the preliminaries on sentiment classification,

language representation, and hybrid quantum-classical machine learning approach in Section II before describing our proposed method in Section III. Next, we present our experimental results and provide relevant discussion in Section IV. Lastly, Section V concludes the paper and describes possible future work.

II. PRELIMINARIES

In this section, we provide the essential theoretical background to aid the readers in understanding our proposed methods. In particular, we elaborate on three aspects: classical sentiment classification, language representation for use in quantum natural language processing, and hybrid quantum-classical machine learning.

A. CLASSICAL SENTIMENT CLASSIFICATION

Sentiment analysis (SA) is a branch of NLP that analyzes the opinion, feeling, or sentiment held by a speaker towards some object or idea communicated through a sentence in natural language, as described in [12]. Generally, the sentiment is labeled positive, neutral, negative, or some value in between. A use-case where sentiment analysis is commonly used is in determining public opinion towards some product or service offered, as explained in [13]. Examples of this include customer reviews for some electronic product or restaurant, film or book reviews, user comments on a social media platform, or the general sentiment of a news article.

Sentiment classification is done on the sentence level, where every sentence has a binary sentiment (positive or negative). In classical machine learning, sentiment analysis can be performed through several methods. One of the simplest methods is the Naïve Bayes algorithm [15]. Other than for sentiment analysis, this algorithm is commonly used for classification tasks on the sentence level that depend on its lexical composition, such as spam detection and political bias classification [13]. The Naïve Bayes algorithm calculates the likelihood of each word appearing in a positive-sentiment or negative-sentiment sentence within the training corpus, then determines the sentiment of a new sentence based on the likelihood of its constituent words [16]. Aside from Naïve Bayes, methods based on artificial neural networks are also commonly employed in

sentiment analysis. Conventionally, this method performs pre-processing on each word in the sentiment into a numerical word embedding form, such as word2vec [17]. The encoding is then used as the input of a neural network. It has been thoroughly reviewed that among neural network models, RNN is most suitable for sentiment analysis [18], [19]. It was shown that RNN has the highest metrics and reliability over several diverse datasets. That is due to the recurrent mechanism inherent in RNNs that propagate information sequentially, which is suitable for processing sentences [19].







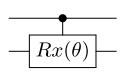
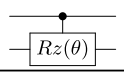
However, neither method accounts for the grammatical structure of the sentence. Naïve Bayes is inherently a bag-of-words model, while RNNs process inputs sequentially in pairs or windows of larger size. Meanwhile, [20] acknowledges that fine-grained sentiment analysis takes grammatical structure into account. Upon considering grammatical structure, [21] discovered that each part of speech has varying significance toward the sentiment of a sentence. The paper ranks each part of speech according to importance or relevance. They identify nouns as having the least importance, while verbs and adjectives as having the highest importance. The distinguishing of part-of-speech contribution may improve model performance within sentiment analysis, as well as provide a possible heuristic for rewriting.

B. LANGUAGE REPRESENTATION FOR QUANTUM NATURAL LANGUAGE PROCESSING

One of the frameworks for natural language processing is the DisCoCat model, derived from the terms distributional, compositional, and categorical. The model stems from [1], which utilizes Lambek calculus and categorical grammar to represent sentences and defines a mapping from pregroup types into a distributional and a numerical in the form of vector space. As such, this model reflects the compositional property of sentences, as the semantic evaluation of a sentence depends on the way its constituent phrases or words are composed. The exact pregroup types belonging to the category may vary. However, most commonly, there exists at least two grammatical types: *n*, which represents the noun type, and *s*, which represents the sentence type. Sentences represented in DisCoCat are depicted using string diagrams, which may be manipulated using a set of graphical calculus rules. As such, it is possible to manipulate and rewrite string diagrams while ensuring the meaning of the sentence is conserved. Interpretation of the diagrams is performed by considering the information propagated from word boxes throughout strings to other word boxes, where the meaning value of a word is evaluated at each word box, and the meaning value of a sentence is evaluated at the open terminating *s* string [1].

DisCoCat is used as the language representation of the QNLP pipeline in [4] and [6]. DisCoCat has been used for question-answering [22], language translation [23], relative pronoun detection [24], and sentiment classification [9], [10], [11].

TABLE 1. List of quantum gates used in the experiment.

Gate Name	Symbol	Matrix
Hadamard		$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$
X		$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
Z		$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
$R_x(\theta)$ (X-rotation)		$\begin{bmatrix} \cos(\frac{\theta}{2}) & -i \sin(\frac{\theta}{2}) \\ -i \sin(\frac{\theta}{2}) & \cos(\frac{\theta}{2}) \end{bmatrix}$
$R_z(\theta)$ (Z-rotation)		$\begin{bmatrix} e^{-i\frac{\theta}{2}} & 0 \\ 0 & e^{i\frac{\theta}{2}} \end{bmatrix}$
$CNOT_{0,1}$		$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$
$CR_{x0,1}(\theta)$		$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos(\frac{\theta}{2}) & -i \sin(\frac{\theta}{2}) \\ 0 & 0 & -i \sin(\frac{\theta}{2}) & \cos(\frac{\theta}{2}) \end{bmatrix}$
$CR_{z0,1}(\theta)$		$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{-i\frac{\theta}{2}} & 0 \\ 0 & 0 & 0 & e^{i\frac{\theta}{2}} \end{bmatrix}$

C. QUANTUM CIRCUITS

A universal quantum computer (UQC), also known as a quantum Turing machine (QTM), is a quantum-mechanics-based computational network consisting of connected qubits. The foundations of this paradigm lie in the universal Turing machine concept that is applied in modern conventional computers. In a universal quantum computer, qubits are assembled in a quantum circuit, where circuit wires propagate information in the form of quantum states, and states are transformed or manipulated using quantum gates. The resulting quantum states of the circuit are then measured in order to determine the result of the intended calculation [25], [26].

A gate is a matrix operation performed on a qubit to transform the state of a quantum circuit. A defining property of a qubit gate is its reversibility, i.e., there exists some other matrix operation that is capable of returning the state of a quantum circuit to its input state after undergoing some qubit gates [25].

In this paper, quantum circuits are composed using an array of quantum gates. In particular, we employ the Hadamard gate (*H*), Pauli-X gate (*X*), Pauli-Z gate (*Z*), X-rotation gate (R_x), Z-rotation gate (R_z), controlled Pauli-X gate or controlled-NOT gate (*CNOT*), controlled X-rotation gate (CR_x), and controlled Z-rotation gate (CR_z). Their definitions stem from [27] reconciled with discopy’s quantum gate implementations. The details of each gate’s symbol and the

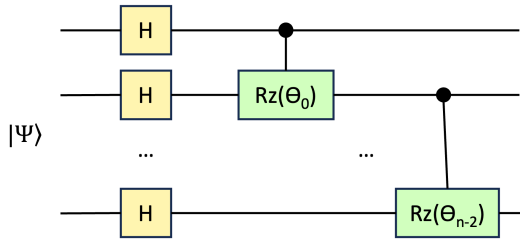


FIGURE 3. Generic form of IQP ansatz.

corresponding matrix operation are listed in Section II-C. Note that the controlled gate definition assumes that for a two-qubit state $|\Psi\rangle \otimes |\Phi\rangle$, $|\Psi\rangle$ is the control qubit and $|\Phi\rangle$ is the target qubit, which would be the convention when defining the controlled operation, e.g., controlled Pauli-X or CNOT, controlled X-rotation, and controlled Z-rotation gates.

D. ANSATZ

In physics and engineering, an ansatz is understood as an initial guess or estimation for some value of interest, to be subsequently adjusted such that the value best approaches the true value [28]. In quantum physics, modeling the wavefunction of a quantum system often uses an ansatz as an initial guess, commonly as a trial wavefunction. Fig. 3 shows an example of a generic form of IQP ansatz. The circuit consists of n qubits appended with H gates, and each qubit is entangled to the subsequent qubit by a series of CR_z gates.

In the context of quantum-circuit-based machine learning, an ansatz is a collection of subroutines or algorithms to construct some initial circuit from a reference form or template, where the resulting circuit possesses a parameterized (trainable) section [29]. Ansatz regulates how quantum circuits are generated from the problem model and determines, for example, how quantum gates are placed within the circuit, the entanglement between qubits, the depth of the circuit, what modifications are made during post-processing, and other aspects of the circuit that may be controlled during the training process.

E. HYBRID QUANTUM-CLASSICAL MACHINE LEARNING

Variational quantum algorithms or VQAs are a class of hybrid quantum-classical machine learning algorithms which employ quantum computers to evaluate quantum circuits, then optimize the circuit parameters on a classical computer. As an optimization method, the VQA task is formally to find the values of θ , such that $\frac{\partial C(\theta)}{\partial \theta} = 0$ for some objective (cost) function $C(\theta)$. For some initial input state ρ , a set of observables O , and some circuit unitary operator $U(\theta)$, the objective function can be generally described as $C(\theta) = f(\rho, O, U(\theta))$. The objective function may also be written as a sum of objective scores at each observation point as in Eq. (1) [3].

$$C(\theta) = \sum_k f_k(\text{Tr}[O_k U(\theta) \rho_k U^\dagger(\theta)]) \quad (1)$$

In designing a VQA, some design choices must be made regarding the VQA components used. Firstly, an ansatz is needed to convert the problem from a descriptive form into a quantum representation. Ansätze are subroutines or initial templates to build an estimate circuit model such that learning can be done to obtain an optimal solution [8]. Within the universal quantum computer paradigm, VQA ansätze have a general form as in Eq. (2) [3].

$$U(\theta) = U_L(\theta_L) \dots U_2(\theta_2) U_1(\theta_1) \quad (2)$$

$$U_l(\theta_l) = \prod_m e^{i\theta_m H_m} W_m \quad (3)$$

where $U_l(\theta_l)$ in Eq. (3) is the unitary at the l -th element of θ , H_m is some Hermitian operator, and W_m is some non-parameterized gate comprising the non-trainable of fixed component of the ansatz.

As VQA defines the model feature space by the quantum circuits used, then the ansatz choice will directly influence the model’s capability to search for the best solution in the hypothesis space. The ansatz is expected to represent the problem sufficiently well. Therefore, ansatz choice is a critical part of designing a machine learning model using the VQA framework [3].

Historically, Instantaneous Quantum Polytime (IQP) ansatz is used for QNLP. IQP is a broad class of quantum circuits that consists of a fixed component and a parameterized component. The circuit is described as instantaneous, that is, the evaluation of quantum gates within the circuit does not depend on the propagation time of information along the qubit wires. The circuit is also described as polytime, meaning that the process of circuit evaluation consumes at most a polynomial number of resources [30]. In QNLP, it was implemented as a layered block construction consisting of a row of Hadamard gates interleaved with controlled unitary rotation gates, and first presented in [6].

This ansatz is also used in [9]. In determining the sentiment of a sentence, [9] expands the pipeline offered in [6] by analyzing the quantum representation suitable for sentiment sentences. It models negation word boxes, i.e., the word “not”, as a fixed quantum gate. The work explores two alternative placements, on the word box and rewritten at the output string, as well as two alternative gate constructions, as a Pauli-X gate and as a Pauli-X gate followed by a Pauli-Z gate. The work also designs auxiliary algorithms for sentiment sentences to conform to the bigraph rewriting method commonly employed.

III. PROPOSED ANSATZ – SIMPLE SENTIMENT ANALYSIS ANSATZ (SIMPLESA)

This paper designs an alternative ansatz to IQP that is specifically explored for sentiment classification. The design motivation stems from a generalization of the IQP layered structure. An essential difference is the observation in [21] that nouns have the least relevance in sentiment analysis, and

TABLE 2. Types of sentences within the dataset.

Type	Component	Number in training data	Number in testing data	Example
SVO	Subject + Verb + Object	7	3	I loved the restaurant.
SVAO	Subject + Verb + Adjective + Object	107	40	I disliked the rude staff.
SXA	Subject + "to be" + Adjective	56	17	The meal was delicious.

as such a strategy may be employed where the impact of noun word boxes in the PQC's is adjusted accordingly.

A. DATA DESCRIPTION

The dataset of sentiment sentences used is the same as in [9](https://github.com/abiwardani/qnlp-sa-fp). The dataset consists of 270 English sentences, with a distribution of 160 training sentences, 50 validation sentences, and 60 testing sentences. The vocabulary is comprised of 29 words: 6 nouns, 8 verbs, and 14 adjectives. The detail of sentence types is presented in Table 2.

Within the training and testing data, there are 28 and 9 negation sentences respectively. All negation sentences are of the form SX[not]A, that is, the word "not" is attached to the adjective (e.g., "the food was not great"). Additionally, the longest sentences in the dataset are those with SVAO and SX[not]A form at 5 words.

The sentences are labeled 0 or 1 indicating negative and positive sentiment respectively. As data is loaded onto the runtime environment, the labels must be mapped onto a data type that is compatible with quantum circuit outputs. As output strings of the diagrams will later be instantiated as one qubit, the data is therefore mapped onto $|0\rangle$ and $|1\rangle$ single qubit states. As a convention, positive (1) labels are mapped to $|0\rangle$ or $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, whereas negative labels are mapped to $|1\rangle$ or $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

B. QNLP PIPELINE

The QNLP pipeline implemented in this research is adapted from the pipelines proposed in [6] and [9]. The abbreviated QNLP pipeline is shown in Fig. 4. There are five main stages: sentence-to-diagram conversion, diagram-to-circuit conversion, optimization, and metric measurement.

The complete pipeline is shown in Fig. 5. Initially, the dataset is loaded onto Python. The dataset is divided into several sets; training, development, and testing sets. As NOT-boxes will be applied, a shortcut for rewriting (as done in [9]) is performed, whereby the word "not" in sentences is briefly ignored, to be reintroduced at a later stage. The sentences are then parsed using Lambeq's Bobcat Parser to obtain raw string diagrams. These diagrams are first filtered based on sentence validity. As explained in [1], the valid sentence is one where its Preg representation reduces to a singular s (sentence) type. These filtered diagrams are then rewritten using Lambeq's default rewrite rules, which are a set of grammatical rewrite rules that perform

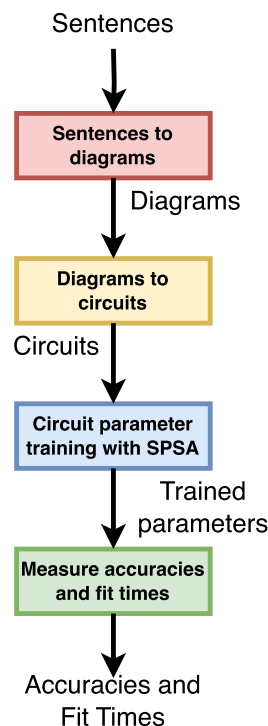


FIGURE 4. General QNLP pipeline for sentiment classification.

logical word box modeling. The resulting diagrams are subsequently transformed into the normal form using the trailing cups conversion algorithm proposed in [9]. This allows the string diagrams to be converted into bigraph form. This method removes cups from the diagrams, thereby resulting in diagrams with fewer qubit wire requirements. The diagrams then undergo stemming, which reconciles conjugated or modified words that have the same stems. The stemmed diagrams are now ready for circuit instantiation.

Step 7 (ansatz application) requires an ansatz to first be chosen. This step differs most greatly from [4], [6], and [9] as both papers use IQP ansatz. After the ansatz is applied and the quantum circuits are generated, the NOT-box is reapplied to negated sentences, following the method outlined in [9]. Where the initial sentence has the word "not" in it, the Pauli-X gate is attached to the output qubit wire. The resulting circuits are now ready to be trained.

To train the circuits using SPSA, the necessary training functions must be compiled. SPSA defines a prediction function and a cost function, where the cost function utilizes the prediction function. These functions call the quantum backend with a set amount of shots, and the resulting

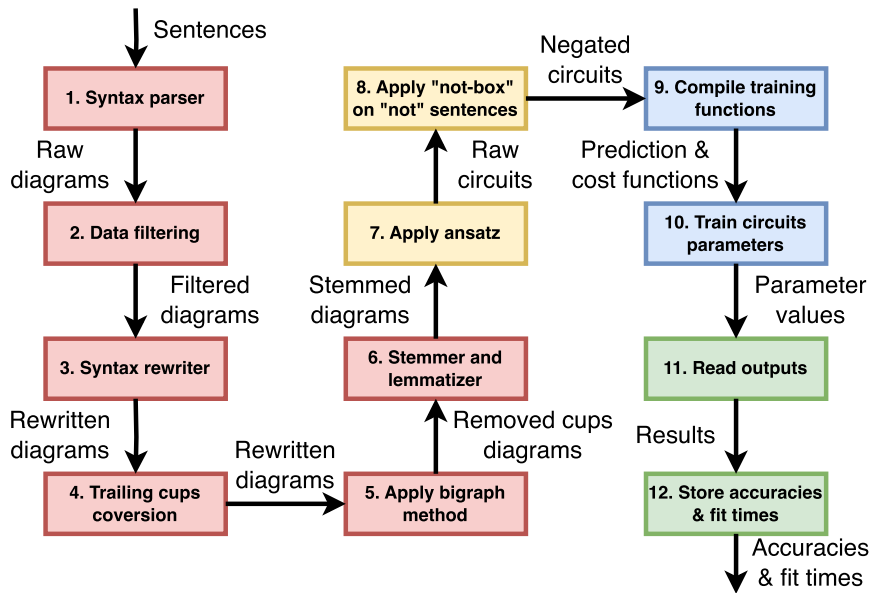


FIGURE 5. Complete QNLP pipeline for sentiment classification.

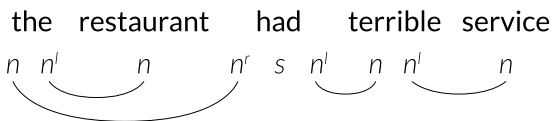


FIGURE 6. Pregroup grammar representation of the sentence “the restaurant had terrible service.”

measurements are recorded. The prediction function takes the average output state, whereas the cost function takes the binary cross-entropy loss against the data labels. The circuit parameters are then trained using an SPSA optimizer. The outputs of the test circuits are read, and the accuracies and fit times are recorded.

A complete example of the pipeline is given as follows. First, take the sample sentence “the restaurant had terrible service” with a negative sentiment label (0). The sentence is loaded onto the Python runtime, and the label is converted into a $|1\rangle$ vector. The sentence is then parsed to obtain the pregroup grammar representation depicted in Fig. 6.

The syntax tree can be viewed as a string diagram consisting of word boxes and wires, as shown in Fig. 7. This string diagram is now ready for rewriting. Rewriting is performed on the string diagram, using the determiner rewrite rule, which treats the determiner word box “the” as a $n \otimes n^l$ cap. This yields the diagram in Fig. 8. The arrangement of cups and caps may be removed using the yanking equations. This normalizes the diagram, resulting in the string diagram depicted in Fig. 9

Then, the cups present in the diagram are converted into trailing form. This form is an adaptation used in [9] to conform to the Python-specific implementation of the bigraph method algorithm. The trailing cups form orders the cups in a cascading manner, such that scanning the diagram top-to-bottom encounters the rightmost cup first and

consecutively encounters cups to the left until it reaches the leftmost cup last. A string diagram depiction of the trailing cups form can be viewed in Fig. 10

The string diagram can now be converted into bigraph form using the bigraph method. Word boxes at even distances from the root word box remain as states, while word boxes at odd distances are transposed into effects. The result is depicted in Fig. 11.

Finally, the string diagram is ready for ansatz consumption and transcribed as a PQC. In this example, IQP ansatz is used. As such, single qubit word boxes (in this case, nouns) are instantiated using XZX Euler decomposition, whereas compound-type word boxes are instantiated as a layered set of Hadamard gates with a linearly cascading set of controlled Z-rotation gates. A compound-type word box is defined as a word box whose domain is more than one *Preg* type. An example is the “had” and “terrible” boxes (outlined in dotted lines) in Fig. 12. The resulting PQC is depicted in Fig. 12.

The process is repeated for all sentences in the dataset. The training and development sets are fed into the SPSA optimizer to learn the rotation gate parameter values. The time required to fit the model onto the training and development sets is recorded. Once the parameters are learned, the accuracy of the test set is measured and recorded. The complete QNLP pipeline is implemented in Python, and structured as a model experimentation notebook which can be accessed at <https://github.com/abiwardani/qnlp-sa-fp>.

C. NOUN PARAMETERIZATION

Drawing from [21] where nouns have the least importance in the evaluation of a sentence’s sentiment, this paper employs an ansatz strategy whereby the impact of nouns is neglected.

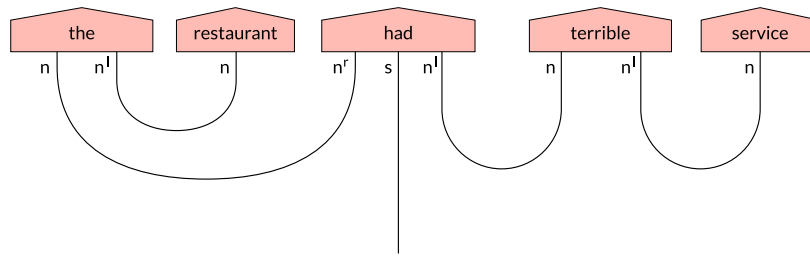


FIGURE 7. String diagram of the sentence "the restaurant had terrible service."

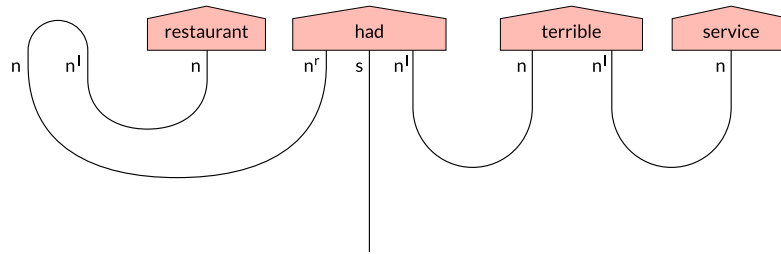


FIGURE 8. Rewritten string diagram of the sentence "the restaurant had terrible service."

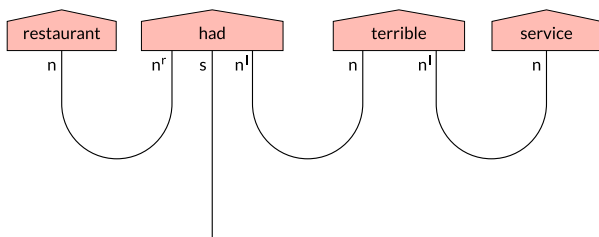


FIGURE 9. Yanked string diagram of the sentence "the restaurant had terrible service."

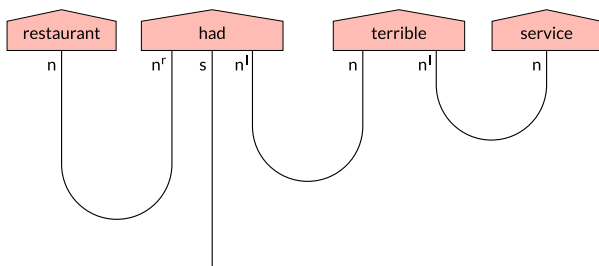


FIGURE 10. Trailing cups form of the sentence "the restaurant had terrible service."

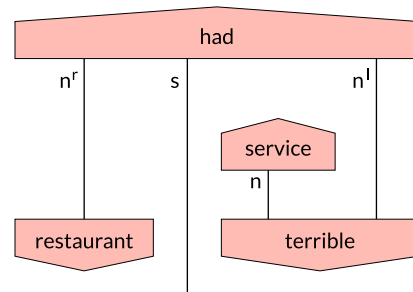


FIGURE 11. Bigraph form of the sentence "the restaurant had terrible service."

It is initially implemented as Lambeq’s IQP ansatz, where nouns are not parameterized. In terms of string diagram rewriting, this may be interpreted as a grammatical rewrite rule which maps n type word boxes as identity wires. In terms of ansatz implementation, this is interpreted as a construction rule which instantiates single n types as the identity gate. This is in contrast to conventional IQP, which defines single n types as an $SU(2)$ gate, instantiated using

XZX Euler decomposition. An example IQP circuit without noun parameterization is depicted in Fig. 13.

However, it is visible from the circuit diagram that this ansatz would result in largely non-predictive circuits. Recalling that controlled Z-rotation gates have matrix form as in Section II-C, the open qubit wire in Fig. 13 would only experience a Hadamard and phase shift.

A concise proof is given in the following. It considers a section of the circuit diagram with two qubits, where the two qubits comprise an equal or larger compound-type word box. The circuit diagram is depicted in Fig. 14. In this, the first qubit starts with a Hadamard gate and has experienced bigraph rewriting, hence ends with a $|0\rangle$ post-selected non-parameterized noun wire, and the second qubit consists of a Hadamard gate followed by a controlled Z-rotation gate. The second qubit then may be used as the control qubit for further sections of the compound-type component, which in turn may experience $|0\rangle$ post-selection.

For this circuit, the final wavefunction is first determined. As a convention for writing this paper, quantum states are

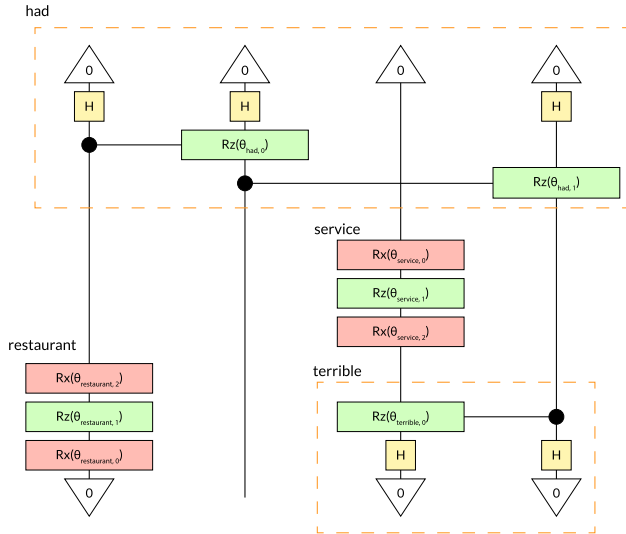


FIGURE 12. IQP circuit of the sentence “the restaurant had terrible service.”

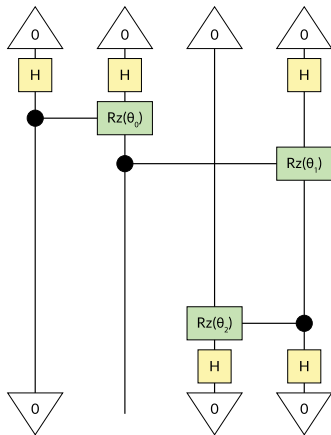


FIGURE 13. IQP circuit diagram without noun parameterization.

written using Big Endian notation. Eq. (4) derives the final wavefunction of the circuit.

$$\begin{aligned}
 |\Psi_0\rangle &= |0\rangle^{\otimes 2} = |00\rangle \\
 |\Psi_1\rangle &= H^{\otimes 2} |\Psi_0\rangle \\
 &= \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle) \\
 |\Psi_2\rangle &= CRz(\theta) |\Psi_1\rangle \\
 &= \frac{1}{2}(|00\rangle + |01\rangle + e^{-i\frac{\theta}{2}} |10\rangle + e^{i\frac{\theta}{2}} |11\rangle) \quad (4)
 \end{aligned}$$

To calculate the desired output probability, *i.e.*, the probability distribution attributed to the second qubit, the calculation must consider the post-selection of the first qubit. The example is when taking the measurement of the $|0\rangle$ state. This can occur in $|00\rangle$ and $|10\rangle$. However, the first qubit is post-selected for $|0\rangle$. As such, its evaluation is interpreted as the conditional probability of measuring $|00\rangle$ given $|00\rangle$ or $|01\rangle$ is measured. Eqs. (5) and (6) calculate the probabilities

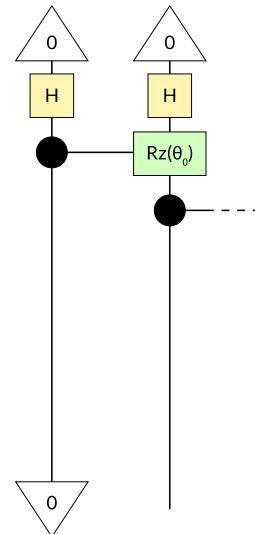


FIGURE 14. Section of IQP circuit without noun parameterization.

of measuring $|00\rangle$ and $|01\rangle$ respectively.

$$\begin{aligned}
 P(|00\rangle) &= |\langle 00|\Psi_2\rangle|^2 \\
 &= |\langle 00|\frac{1}{2}(|00\rangle + |01\rangle + e^{-i\frac{\theta}{2}} |10\rangle + e^{i\frac{\theta}{2}} |11\rangle)|^2 \\
 &= |\langle 00|\frac{1}{2}|00\rangle|^2 = \frac{1}{4} \quad (5)
 \end{aligned}$$

$$\begin{aligned}
 P(|01\rangle) &= |\langle 01|\Psi_2\rangle|^2 \\
 &= |\langle 01|\frac{1}{2}(|00\rangle + |01\rangle + e^{-i\frac{\theta}{2}} |10\rangle + e^{i\frac{\theta}{2}} |11\rangle)|^2 \\
 &= |\langle 01|\frac{1}{2}|01\rangle|^2 = \frac{1}{4} \quad (6)
 \end{aligned}$$

The conditional probability is calculated in Eq. (7). As shown, the qubit has a probability of measuring $|0\rangle$ at 0.5. This calculation can be expanded to the four-qubit circuit shown in Fig. 13. Notice that the transposed circuit section comprising the compound-type block and $|0\rangle$ post-selections on qubits 3 and 4 is a rotated copy of the 2-qubit base case. Likewise, the controlled Z-rotation with control qubit 2 and target qubit 4 only experiences phase shift, and as such does not meaningfully post-select the second qubit in the computational basis.

$$\begin{aligned}
 P(O = |0\rangle) &= P(|0\rangle_2 | |0\rangle_1) \\
 &= \frac{P(|00\rangle)}{P(|00\rangle) + P(|01\rangle)} \\
 &= \frac{1/4}{1/4 + 1/4} = \frac{1}{2} \quad (7)
 \end{aligned}$$

Given that this is independent of the value of θ . Thus, it indicates a QNLP model using IQP ansatz without nouns would result in a model that is incapable of learning. Furthermore, the model would only infer $|0\rangle$ and $|1\rangle$ sentiments at a uniform rate. As such, the model would be non-predictive. An alternative compound-type construction must be determined.

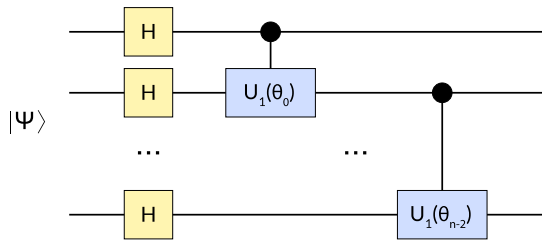


FIGURE 15. SimpleSA Generic Compound-Type.

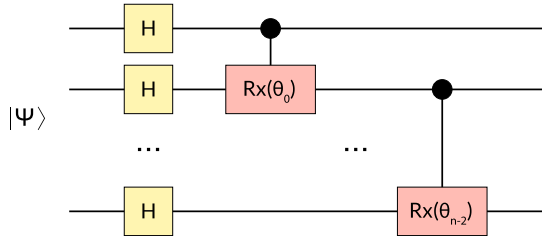


FIGURE 16. SimpleSA H-CRx compound-type construction.

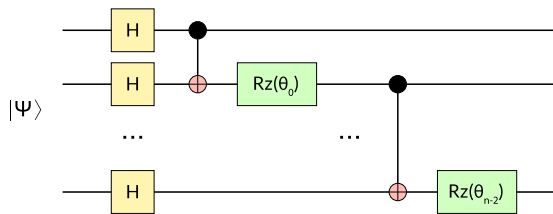


FIGURE 17. SimpleSA H-CNOT-Rz compound-type construction.

D. COMPOUND-TYPE CONSTRUCTION

Following from Chapter III-C, a cascading controlled Z-rotation construction is not suitable for a noun non-parameterization strategy. The compound-type construction must hence be replaced. The template however still follows the linearly-entangled cascading structure of Lambeq’s implemented IQP ansatz. This alternative ansatz is based on IQP, and it employs noun non-parameterization. As it is designed for sentiment classification, this paper refers to it as the Simple Sentiment Analysis ansatz, or SimpleSA ansatz. The generic SimpleSA compound-type construction, therefore, resembles Fig. 15.

The following three constructions are investigated in this paper: H-CRx, H-CNOT-Rz, and H-CNOT-Rz-H. The H-CRx construction is very similar to IQP ansatz, where the CRz gates are swapped for CRx gates. A diagram depicting the H-CRx compound-type construction is shown in Fig. 16. The H-CNOT-Rz construction has a cascading layer of CNOT gates, where each CNOT gate is followed by a Z-rotation gate. The H-CNOT-Rz compound-type construction is depicted in Fig. 17. The H-CNOT-Rz-H construction is similar to the H-CNOT-Rz construction but has an additional Hadamard gate following the Z-rotation gate. The H-CNOT-Rz-H compound-type construction is depicted in Fig. 18. Sample circuits for each construction are depicted in Figs. 19, 20, and 21. Notice that the set of controlled gates follow the patterns shown in the compound-type construction diagrams.

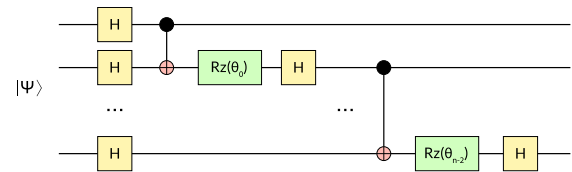


FIGURE 18. SimpleSA H-CNOT-Rz-H compound-type construction.

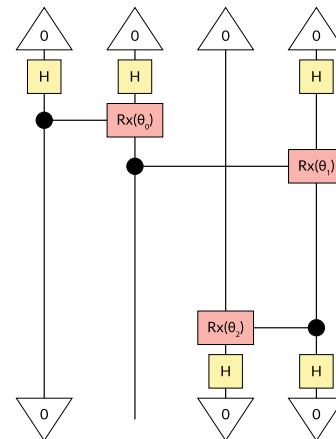


FIGURE 19. SimpleSA H-CRx circuit diagram.

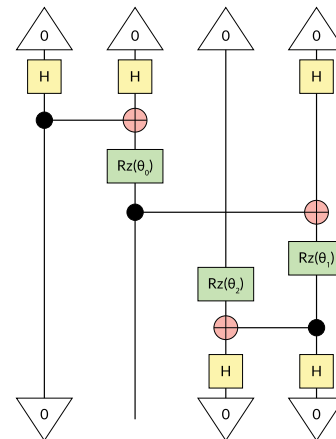


FIGURE 20. SimpleSA H-CNOT-Rz circuit diagram.

Algorithm 1 describes a code for constructing SimpleSA circuit segments using H-CNOT-Rz-H block components. The code is intended to be used modularly, taking sections of the string diagram based on a word box, and then composing the resulting circuits together to form a final circuit. Upon processing each segment, there are two types of word boxes that may occur, as previously explained: single qubit types and compound-types. The procedure takes as input the number of qubits spanned by the word box and a list of variables *params*. If the number of qubits is 1, then the word box is a single qubit type. For the scope of this paper, single qubit types are noun word boxes. In this case, nouns may be parameterized or not parameterized. As such, the pseudocode checks the length of *params*; if it

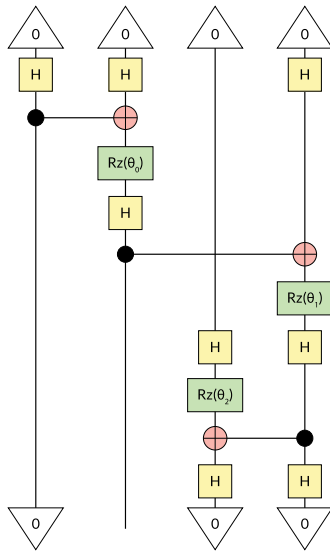


FIGURE 21. SimpleSA H-CNOT-Rz-H circuit diagram.

is zero, then the noun is not parameterized, and the circuit is instantiated as a single qubit identity wire. Otherwise, the noun is parameterized, and the circuit is instantiated as an Rx-Rz-Rx component with the first, second, and third elements of *params* as its gate parameters.

On the other hand, if the number of qubits spanned by the word box is more than 1, then the word box is of a compound-type. Therefore, the cascading entangling gates are used. This is built starting with a row of Hadamard gates spanning the qubit register, followed by a series of CNOT then Rz gates parameterized as the *i*-th element of *params*, where *i* is the index of the qubit register. The Hadamards and rotation components are repeated with the depth of *params*.

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

In this section, we present some experiments to highlight our proposed method, Simple Sentiment Analysis Ansatz (SimpleSA). The proposed ansatz requires no noun parameterization with a smaller amount of quantum gates, but it performs better than other ansatz. The dataset used for the experiments is the same as explained in Section III-A. As an environmental simulation setup, we use an 8-core vCPU with 52 GB RAM in Linux with VM type of n1-highmem-8.

The results of our experiments are presented in Table 3. As shown, SimpleSA Ansatz with default parameters and H-CNOT-CRz-H block construction outperforms baseline IQP Ansatz at 85.00% against 72.50%, displaying a significant 12.50% increase in accuracy. Noun parameterization slightly decreases SimpleSA ansatz performance. Importantly, SimpleSA ansatz with H-CRx and H-CNOT-Rz construction displays non-predictive behavior similar to IQP ansatz without nouns, as indicated by the asterisk (*) in Table 3.

Moreover, Fig. 22 shows the comparison of training accuracy for several ansätze. In this case, we only compare IQP Ansatz, SimpleSA (SSA) Ansatz without noun

Algorithm 1 SimpleSA Circuit With H-CNOT-Rz-H Construction

Input: *nQubits*: Number of qubits used in input word box, *params*: list of parameters used in circuit.

Output: *circuit*: resulting SimpleSA circuit

begin

integer *nQubit*

arraySymbol *params*

Circuit *circuit*

Circuit hadamards, rotations

integer *N* ← length of *params*

if *nQubits* = 1 **then**

if *N* = 0 **then**

circuit ← Id(1)

else

circuit ←

 Rx(*params*₁) ◦ Rz(*params*₂) ◦ Rx(*params*₃)

else

assert *params* is 2-dimensional

assert number of parameters for each qubit equals *nQubits*-1

circuit ← Id(*nQubits*)

for each *thetasBatch* ∈ *params* **do**

hadamards ← $H^{\otimes nQubits}$

rotations ← Id(*nQubits*)

for *i* = 0 to *nQubits* - 2 **do**

rotations ← *rotations* ◦ (Id(*i*) ⊗ CX ⊗ Id(*nQubits* - 2 - *i*))

rotations ← *rotations* ◦ (Id(*i* + 1) ⊗

 Rz(*thetasBatch*_{*i*}) ⊗ Id(*nQubits* - 2 - *i*))

rotations ← *rotations* ◦ (Id(*i* + 1) ⊗

 H ⊗ Id(*nQubits* - 2 - *i*))

circuits ← *circuits* ◦ *hadamards* ◦ *rotations*

return *circuit*

parameterization and SimpleS (SSA) Ansatz with noun parameterization. We run the training for 130 iterations. As shown in Fig. 22, IQP Ansatz with noun parameterization (IQP-Noun), SimpleSA without noun parameterization (SSA), and SimpleSA with noun parameterization (SSA-Noun) have training accuracy around 70-85%. However, the training results are then further tested as shown in Table 3.

The confusion matrices for each SVO, SVAO, and SXA sentence type are presented in Tables 4 and 5 for IQP and SimpleSA model respectively. The columns “Actual |0)” and “Actual |1)” record instances where the test sentence is labelled |0) and |1) respectively. The rows “Predicted |0)” and “Predicted |1)” record instances where the model classifies the test sentence as |0) or |1), respectively.

A. NON-PREDICTIVE ANSÄTZE

From our experiment, it was found that H-CRx and H-CNOT-Rz constructions result in non-predictive SimpleSA

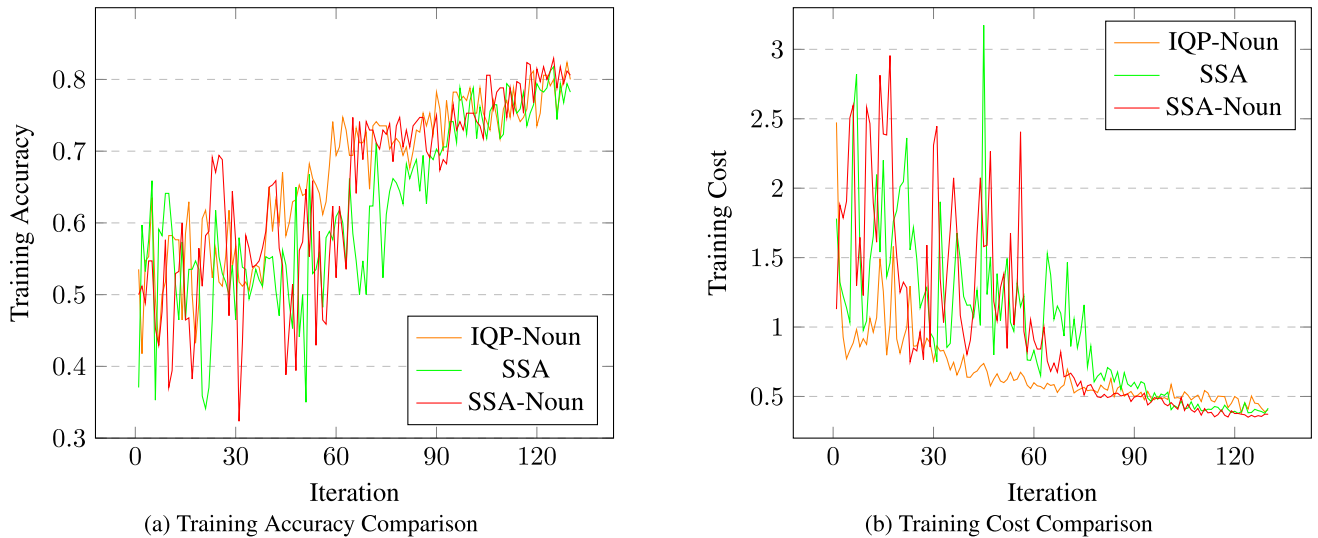


FIGURE 22. Sentiment analysis training accuracy and cost comparisons for various Ansatz.

TABLE 3. Testing accuracies and fit times of experiment results.

Ansatz	Compound-Type Construction	Noun Parameterization	Testing Accuracy	Fit Time
IQP	H-CRz	Yes	72.50%	02:42:18
	H-CRz	No	50.00%*	-
SimpleSA (SSA)	H-CRx	No	50.00%*	-
	H-CNOT-Rz	No	50.00%*	-
	H-CNOT-Rz-H	No	85.00%	02:14:49
		Yes	82.50%	03:23:54

models. Derivation can be performed in a similar manner as in Chapter III-C using Figs. 19 and 20 as reference.

For SimpleSA ansatz with H-CRx construction, its wavefunction can be derived as in Eq. (8)

$$\begin{aligned}
 |\Psi_0\rangle &= |0\rangle^{\otimes 2} = |00\rangle \\
 |\Psi_1\rangle &= H^{\otimes 2} |\Psi_0\rangle = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle) \\
 |\Psi_2\rangle &= CRx(\theta) |\Psi_1\rangle \\
 &= \frac{1}{2}(|00\rangle + |01\rangle + (\cos\left(\frac{\theta}{2}\right) - i \sin\left(\frac{\theta}{2}\right))|10\rangle \\
 &\quad + (\cos\left(\frac{\theta}{2}\right) - i \sin\left(\frac{\theta}{2}\right))|11\rangle) \tag{8}
 \end{aligned}$$

Following the derivation method outlined in Chapter III-C, the probabilities of measuring $|00\rangle$ and $|01\rangle$ are calculated in Eqs. (9) and (10). The probability of measuring $|0\rangle$ on the second qubit is likewise calculated using the conditional probability under post-selection, as derived in Eq. (11). This yields another non-predictive model whose inferences do not depend on circuit parameters and it measures single qubit states of maximum superposition.

$$\begin{aligned}
 P(|00\rangle) &= |\langle 00|\Psi_2\rangle|^2 \\
 &= |\langle 00|\frac{1}{2}(|00\rangle + |01\rangle)
 \end{aligned}$$

$$\begin{aligned}
 &+ (\cos\left(\frac{\theta}{2}\right) - i \sin\left(\frac{\theta}{2}\right))|10\rangle \\
 &+ \cos\left(\frac{\theta}{2}\right) - i \sin\left(\frac{\theta}{2}\right)|11\rangle|^2 \\
 &= |\langle 00|\frac{1}{2}|00\rangle|^2 = \frac{1}{4} \tag{9}
 \end{aligned}$$

$$\begin{aligned}
 P(|01\rangle) &= |\langle 01|\Psi_2\rangle|^2 \\
 &= |\langle 01|\frac{1}{2}(|00\rangle + |01\rangle \\
 &\quad + (\cos\left(\frac{\theta}{2}\right) - i \sin\left(\frac{\theta}{2}\right))|10\rangle \\
 &\quad + \cos\left(\frac{\theta}{2}\right) - i \sin\left(\frac{\theta}{2}\right)|11\rangle)|^2 \\
 &= |\langle 01|\frac{1}{2}|01\rangle|^2 = \frac{1}{4} \tag{10}
 \end{aligned}$$

$$\begin{aligned}
 P(O = |0\rangle) &= \frac{P(|0\rangle_2 | |0\rangle_1)}{P(|00\rangle)} \\
 &= \frac{P(|00\rangle) + P(|01\rangle)}{P(|00\rangle) + P(|01\rangle)} \\
 &= \frac{1/4 + 1/4}{1/4 + 1/4} = \frac{1}{2} \tag{11}
 \end{aligned}$$

The same derivation is done with the H-CNOT-Rz construction. It similarly describes H-CNOT-Rz circuits as non-predictive, hence unsuitable to use as a VQA ansatz. However, the derivation of H-CNOT-Rz-H circuits shows that it is trainable. In practice, it was this exact trainability that

served as the motivation for applying another Hadamard gate. The wavefunction of the base case H-CNOT-Rz-H circuit is shown in Eq. (12).

$$\begin{aligned}
 |\Psi_0\rangle &= |0\rangle^{\otimes 2} = |00\rangle \\
 |\Psi_1\rangle &= H^{\otimes 2} |\Psi_0\rangle = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle) \\
 |\Psi_2\rangle &= CNOT |\Psi_1\rangle \\
 &= \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle) \\
 |\Psi_3\rangle &= (I \otimes Rz(\theta)) |\Psi_2\rangle \\
 &= \frac{1}{2}(e^{-i\frac{\theta}{2}} |00\rangle + e^{i\frac{\theta}{2}} |01\rangle + e^{-i\frac{\theta}{2}} |10\rangle + e^{i\frac{\theta}{2}} |11\rangle) \\
 |\Psi_4\rangle &= H^{\otimes 2} |\Psi_3\rangle \\
 &= \frac{1}{2} \left((e^{-i\frac{\theta}{2}} + e^{i\frac{\theta}{2}}) |00\rangle + (e^{-i\frac{\theta}{2}} - e^{i\frac{\theta}{2}}) |01\rangle \right) \quad (12)
 \end{aligned}$$

Calculating the conditional probabilities uses the probabilities of measuring $|00\rangle$ and $|01\rangle$, which are derived in Eqs. (13) and (14) respectively.

$$\begin{aligned}
 P(|00\rangle) &= |\langle 00 | \Psi_4 \rangle| \\
 &= |\langle 00 | \frac{1}{2} \left((e^{-i\frac{\theta}{2}} + e^{i\frac{\theta}{2}}) |00\rangle + (e^{-i\frac{\theta}{2}} - e^{i\frac{\theta}{2}}) |01\rangle \right) | \\
 &= |\langle 00 | \frac{1}{2} (e^{-i\frac{\theta}{2}} + e^{i\frac{\theta}{2}}) |00\rangle | \\
 &= \frac{1}{4} |e^{-i\frac{\theta}{2}} + e^{i\frac{\theta}{2}}| \\
 &= \frac{1}{4} |2 \cos\left(\frac{\theta}{2}\right)| \\
 &= \cos^2\left(\frac{\theta}{2}\right) \quad (13)
 \end{aligned}$$

$$\begin{aligned}
 P(|01\rangle) &= |\langle 01 | \Psi_4 \rangle| \\
 &= |\langle 01 | \frac{1}{2} \left((e^{-i\frac{\theta}{2}} + e^{i\frac{\theta}{2}}) |00\rangle + (e^{-i\frac{\theta}{2}} - e^{i\frac{\theta}{2}}) |01\rangle \right) | \\
 &= |\langle 01 | \frac{1}{2} (e^{-i\frac{\theta}{2}} - e^{i\frac{\theta}{2}}) |01\rangle | \\
 &= \frac{1}{4} |e^{-i\frac{\theta}{2}} - e^{i\frac{\theta}{2}}| \\
 &= \frac{1}{4} |-2i \sin\left(\frac{\theta}{2}\right)| \\
 &= \sin^2\left(\frac{\theta}{2}\right) \quad (14)
 \end{aligned}$$

The probability of measuring $|0\rangle$ on the second qubit is as calculated in Eq. (15). This yields a value of $\cos^2\left(\frac{\theta}{2}\right)$. In this compound-type construction, the output probability does rely on the θ -parameterized rotation gates. Hence, the SimpleSA ansatz with H-CNOT-Rz-H construction is predictive, which explains its behavior against non-predictive circuits in Table 3.

$$\begin{aligned}
 P(O = |0\rangle) &= P(|0\rangle_2 | |0\rangle_1) \\
 &= \frac{P(|00\rangle)}{P(|00\rangle) + P(|01\rangle)}
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{\cos^2\left(\frac{\theta}{2}\right)}{\cos^2\left(\frac{\theta}{2}\right) + \sin^2\left(\frac{\theta}{2}\right)} \\
 &= \cos^2\left(\frac{\theta}{2}\right) \quad (15)
 \end{aligned}$$

B. ACCURACY ANALYSIS

From Table 3, SimpleSA is capable of outperforming IQP ansatz at 85.00% accuracy. This implementation of the quantum sentiment classification pipeline also outperforms the IQP model evaluated in [9], which achieves a maximum accuracy of 81.67%. It is expected that the non-parameterization of noun types allows SimpleSA to evaluate the sentiment of sentences more suitably by focusing on sentiment-relevant words, such as verbs and adjectives [21]. This underperformance of IQP is observable in Tables 4 and 5, where IQP has 58.82% accuracy at SXA classification and SimpleSA achieves 100.00%. This is reflected in the respective expressions for output qubit probability.

First, consider the behavior of IQP models on some generic SXA sentences, shown in Fig. 23. The circuit is generated from two-word boxes: the single qubit noun type instantiated as XZX diagonalization in the bottom part of the first qubit, and the two-qubit adjective spanning the first and second qubits instantiated as a row of Hadamard gates and a controlled Z-rotation gate.

Analytically deriving the form of its final wavefunction and then taking the conditional probability of the output qubit where the first qubit is in the $|0\rangle$ state yields the expression in Eq. (16).

$$P(|0\rangle_2) = \frac{0.5 (\sin(\theta_2) \sin(\theta_3) + 1)}{Q(\theta_0, \theta_1, \theta_2, \theta_3)} \quad (16)$$

where $Q(\theta_0, \theta_1, \theta_2, \theta_3)$ is defined in Eq. (17), θ_0 is the gate rotation parameter of the block component representing the adjective, and $\theta_1, \theta_2, \theta_3$ are the single qubit parameters representing the subject noun in an SXA type sentence.

$$\begin{aligned}
 Q(\theta_0, \theta_1, \theta_2, \theta_3) &= 0.5 \sin(\theta_0) \sin(\theta_1) \cos(\theta_3) \\
 &\quad + 0.5 \sin(\theta_0) \sin(\theta_3) \cos(\theta_1) \cos(\theta_2) \\
 &\quad + 0.5 \sin(\theta_2) \sin(\theta_3) \cos(\theta_0) \\
 &\quad + 0.5 \sin(\theta_2) \sin(\theta_3) + 1.0 \quad (17)
 \end{aligned}$$

It is obvious that the expression for $P(|0\rangle_2)$ depends on the subject noun thetas. Taking a sample SXA sentence “(the) meal was good” and plugging in the learned values of $\theta_1, \theta_2, \theta_3$ for the noun “meal” into Eq. (16), the expression in Eq. (18) is obtained.

$$P(|0\rangle_2) = \frac{0.6823}{0.0398 \sin(\theta_0) + 0.1823 \cos(\theta_0) + 1.1823} \quad (18)$$

Upon differentiating with respect to θ_0 and equating the expression to 0, the range $P(|0\rangle_2) \in [0.4984, 0.6853]$ is obtained. Therefore, the model circuit by nature tends to infer $|0\rangle$ more than $|1\rangle$. This imbalanced output space is undesirable, as there is no linguistic reason that SXA sentences with the subject “meal” should tend to be a

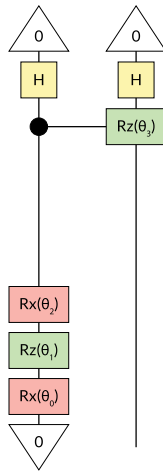


FIGURE 23. IQP circuit of SXA sentence.

TABLE 4. IQP model confusion matrix for each sentence type.

		Actual $ 0\rangle$	Actual $ 1\rangle$
SVO	Predicted $ 0\rangle$	0	1
	Predicted $ 1\rangle$	1	1
SVAO	Predicted $ 0\rangle$	17	5
	Predicted $ 1\rangle$	2	16
SXA	Predicted $ 0\rangle$	5	3
	Predicted $ 1\rangle$	4	5

TABLE 5. SimpleSA model confusion matrix for each sentence type.

		Actual $ 0\rangle$	Actual $ 1\rangle$
SVO	Predicted $ 0\rangle$	1	0
	Predicted $ 1\rangle$	0	2
SVAO	Predicted $ 0\rangle$	12	0
	Predicted $ 1\rangle$	7	21
SXA	Predicted $ 0\rangle$	9	0
	Predicted $ 1\rangle$	0	8

positive sentiment more than a negative one. In contrast, the expression for SimpleSA circuit output probabilities will resemble Eq. (15), and solely focuses on the parameter associated with the adjective. Hence, SimpleSA accurately represents the sentiment analysis aspect of the task, in terms of SXA sentence classification.

C. COMPLEXITY ANALYSIS

This chapter compares the circuit complexity between the three best ansatz configurations: IQP with nouns, SimpleSA without nouns and H-CNOT-Rz-H block construction, and SimpleSA with nouns and H-CNOT-Rz-H block construction. The three models are referred to as IQP-Noun, SimpleSA, and SimpleSA-Noun respectively. Table 6 describes the number of parameters and the number of gates used by each ansätze. As the number of parameters varies for each sentence and experiment setting, its record is written as a function of n, i, w . The variable n denotes the number of noun word boxes present in the string diagram, i denotes the index of compound-type word box in the string diagram, and w_i is the width of the i -th compound-type word box. The parameter counts are derived from each ansatz definition, assuming the

sentence types covered in Chapter III-A. The derivation also assumes that the ansätze are configured for one layer (layers are not repeated). As in both ansätze, multi-qubit word boxes are implemented using linear entanglement and both n and s strings are instantiated as one qubit, therefore the number of parameters in each word box is equal to one less than the number of strings comprising the compound-type word box (denoted w). Additionally, as IQP-Noun parameterizes nouns using three rotations, it follows that IQP parameter count has a $3n$ term. Therefore, in terms of parameter count, SimpleSA (without nouns) uses $3n$ less parameters than IQP-Noun.

In terms of number of gate operations, SimpleSA with H-CNOT-Rz-H construction uses more gates than IQP (which uses H-CRz construction). Specifically, it uses an additional two gates for each controlled gate, i.e. the target qubit wire of the controlled gate is followed by two gates per layer. This contributes to the $2(w_i - 1)$ term of each compound-type word box. Furthermore, all ansätze compound-types initialize with a row of Hadamard gates, hence have an additional w_i term. The resulting gate counts are shown in Table 6.

For our dataset, the total number of parameters present in each set of training circuits are calculated as follows. In *Preg* representation, SVO circuits are reduced to $n^r sn^l$. Likewise, SVAO sentences are reduced to $n^r s(n^l)^l n^l$, and SXA sentences are reduced to $n^r s$. Reduction is performed using the method outlined in Chapter III-B. Using the formulae in Table 6, the values in Table 7 are obtained. Multiplying each number of parameters with the number of each sentence type in Table 2, there are 1243 total parametric gates in IQP-Noun training circuits, 391 total parametric gates in SimpleSA training circuits, and 1243 total parametric gates in SimpleSA-Noun training circuits. In other words, SimpleSA uses about one-third the amount of parametric gates as IQP-Noun and SimpleSA-Noun. Hence, SimpleSA performs about one-third the amount of rotation gate operations as noun-parameterized ansätze. This reduction of rotation gates leads to shorter inference times, which is reflected in the fit times recorded in Table 3.

The total number of gate operations is derived similarly to the number of parametric gates. Table 8 shows the number of gates by sentence type for different ansätze. The following values are obtained: 1911 gate operations for IQP-Noun, 1841 gate operations for SimpleSA, and 2693 gate operations for SimpleSA-Noun. The number of gate operations respectively have high correlation with fit times ($R=0.9454$), indicating that circuit depth increases fit times. This is likely due to increased inference times, where more gate operations leads to longer circuit evaluation. Thus, SimpleSA-Noun takes the longest to train at 3 hours 24.9 minutes while SimpleSA is fastest at 2 hours 14.2 minutes.

In terms of parameter optimization, this also affects the number of unique trainable parameters. By counting the unique parameter symbols over all sentences, it is obtained that IQP-Noun uses 60 unique parameters whereas SimpleSA and SimpleSA-Noun uses 42 unique parameters. This 30% reduction in number of trainable parameters contributes

TABLE 6. Complexity formula for different ansätze.

Model	No. of params	No. of gates
IQP-Noun	$3n + \sum_i (w_i - 1)$	$3n + \sum_i (2w_i - 1)$
SimpleSA	$\sum_i (w_i - 1)$	$\sum_i (w_i + 3(w_i - 1))$
SimpleSA-Noun	$3n + \sum_i (w_i - 1)$	$3n + \sum_i (w_i + 3(w_i - 1))$

TABLE 7. Number of parameters by sentence type of different ansätze.

Model	No. of params			Average
	SVO	SVAO	SXA	
IQP-Noun	8	9	4	7
SimpleSA	2	3	1	2
SimpleSA-Noun	8	9	4	7

TABLE 8. Number of gates by sentence type of different ansätze.

Model	No. of gates			Average
	SVO	SVAO	SXA	
IQP-Noun	11	14	6	10.33
SimpleSA	9	14	5	9.33
SimpleSA-Noun	15	20	8	14.33

to model stability. Where IQP-Noun experiences overfit (achieves 80.00% training accuracy in Fig. 22a and 72.50% testing accuracy in Table 3), SimpleSA and SimpleSA-Noun do not. Most notably, SimpleSA achieves 78.23% training accuracy in Fig. 22a and 85.00% testing accuracy in Table 3. This is due to the reduction of trainable parameters in SimpleSA, which leads to a suitable reduction of hypothesis space, as evidenced by the elimination of overfit present in IQP. For this reason, the non-parameterization of nouns in SimpleSA ansatz is essential to its robustness over IQP-Noun ansatz.

D. COMPARISON WITH OTHER METHODS

Additionally, we also compare our results with previous works in literature (i.e., [9], [10]). In particular, we compare the best-performing model from [9] and our best result. As shown, previous works obtain the maximum accuracy of 81.67% [9]. Meanwhile, ours yields 85% by leaving out noun parameterization, which shows the superiority of our method. Note that another work on sentiment analysis was performed by Ganguly et al., [10], which yields perfect accuracy and 83.33% for noiseless and noisy quantum computation [10], respectively. However, the work utilizes much smaller samples (i.e., a total of 130 sentences: 130, 30, and 30 for training, development, and testing, respectively) with little information on the technical implementation and the sentences employed than [9], which uses a total of 280 samples: 170 for training, 50 for development, and 60 for testing), and this work, whose dataset is as previously shown in Table 2.

V. CONCLUSION AND FUTURE WORK

This paper explored a novel implementation of sentiment analysis using the Variational Quantum Algorithms (VQA) framework. This paper proposed an alternative ansatz for

the sentiment classification task in quantum representation. The proposed SimpleSA has less complexity than the other ansätze in terms of number of parameters and number of gates. Moreover, experimental results showed that the SimpleSA ansatz with H-CNOT-Rz-H compound block construction outperforms IQP ansatz at 85.00% accuracy. Furthermore, SimpleSA optimization converges 20.89% faster than IQP for the SPSA method with 130 iterations.

In conclusion, SimpleSA ansatz is more suitable than IQP ansatz at sentiment classification. It has a circuit construction that is robust to noun choice and suitably determines the sentiment of a sentence based on relevant word boxes. This is due to the employment of a part-of-speech-oriented strategy where noun word boxes are effectively ignored and instantiated as identity wires.

Future work may be done by employing more specific tuning strategies. For example, nouns may be instantiated by rotation gates that do not diverge far from the identity operation. Additionally, other parts of speech may be modeled accordingly to their behavior within sentiment analysis according to existing literature.

ACKNOWLEDGMENT

Muhammad Rifat Abiwardani thanks Agung Budiyo, Ph.D., for the valuable discussions. Rahmat Mulyawan also thanks the Research Collaboration Center for Quantum Technology 2.0, Bandung Institute of Technology.

REFERENCES

- [1] B. Coecke, M. Sadrzadeh, and S. Clark, "Mathematical foundations for a compositional distributional model of meaning," 2010, *arXiv:1003.4394*.
- [2] J. Preskill, "Quantum computing in the NISQ era and beyond," *Quantum*, vol. 2, p. 79, Aug. 2018, doi: [10.22331/q-2018-08-06-79](https://doi.org/10.22331/q-2018-08-06-79).
- [3] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, "Variational quantum algorithms," *Nature Rev. Phys.*, vol. 3, no. 9, pp. 625–644, Aug. 2021, doi: [10.1038/s42254-021-00348-9](https://doi.org/10.1038/s42254-021-00348-9).
- [4] K. Meichanetzidis, S. Gogioso, G. de Felice, N. Chiappori, A. Toumi, and B. Coecke, "Quantum natural language processing on near-term quantum computers," *Electron. Proc. Theor. Comput. Sci.*, vol. 340, pp. 213–229, Sep. 2021, doi: [10.4204/eptcs.340.11](https://doi.org/10.4204/eptcs.340.11).
- [5] W. Zeng and B. Coecke, "Quantum algorithms for compositional natural language processing," 2016, *arXiv:1608.01406*.
- [6] D. Kartsaklis, I. Fan, R. Yeung, A. Pearson, R. Lorenz, A. Toumi, G. de Felice, K. Meichanetzidis, S. Clark, and B. Coecke, "Lambeq: An efficient high-level Python library for quantum NLP," 2021, *arXiv:2110.04236*.
- [7] D. Kartsaklis, M. Sadrzadeh, S. Pulman, and B. Coecke, "Reasoning about meaning in natural language with compact closed categories and Frobenius algebras," in *Logic and Algebraic Structures in Quantum Computing* (Lecture Notes in Logic). Cambridge, U.K.: Cambridge Univ. Press, 2016, pp. 199–222.
- [8] J. Romero, J. P. Olson, and A. Aspuru-Guzik, "Quantum autoencoders for efficient compression of quantum data," *Quantum Sci. Technol.*, vol. 2, no. 4, Aug. 2017, Art. no. 045001, doi: [10.1088/2058-9565/aa8072](https://doi.org/10.1088/2058-9565/aa8072).
- [9] F. Z. Ruskanda, M. R. Abiwardani, M. A. Al Bari, K. A. Bagaspati, R. Mulyawan, I. Syafalni, and H. T. Larasati, "Quantum representation for sentiment classification," in *Proc. IEEE Int. Conf. Quantum Comput. Eng. (QCE)*, Sep. 2022, pp. 67–78.
- [10] S. Ganguly, S. N. Morapakula, and L. M. P. Coronado, "Quantum natural language processing based sentiment analysis using Lambeq toolkit," in *Proc. 2nd Int. Conf. Power, Control Comput. Technol. (ICPC2T)*, Mar. 2022, pp. 1–6.

- [11] J. Stein, I. Christ, N. Kraus, M. B. Mansky, R. Müller, and C. Linnhoff-Popien, "Applying QNLP to sentiment analysis in finance," 2023, *arXiv:2307.11788*.
- [12] B. Liu, *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions* (Studies in Natural Language Processing), 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2020.
- [13] S. Tedmori and A. Awajan, "Sentiment analysis main tasks and applications: A survey," *J. Inf. Process. Syst.*, vol. 15, no. 3, pp. 500–519, 2019.
- [14] V. Martinez and G. Leroy-Meline, "A multiclass Q-NLP sentiment analysis experiment using DisCoCat," 2022, *arXiv:2209.03152*.
- [15] D. Jurafsky and J. H. Martin, *Naive Bayes and Sentiment Classification*, 3rd ed. London, U.K.: Pearson, 2023, pp. 58–79.
- [16] Q. Li, H. Peng, J. Li, C. Xia, R. Yang, L. Sun, P. S. Yu, and L. He, "A survey on text classification: From traditional to deep learning," *ACM Trans. Intell. Syst. Technol.*, vol. 13, no. 2, pp. 1–41, Apr. 2022, doi: [10.1145/3495162](https://doi.org/10.1145/3495162).
- [17] J. Bhatta, D. Shrestha, S. Nepal, S. Pandey, and S. Koirala, "Efficient estimation of Nepali word representations in vector space," *J. Innov. Eng. Educ.*, vol. 3, no. 1, pp. 71–77, Mar. 2020.
- [18] N. C. Dang, M. N. Moreno-García, and F. D. la Prieta, "Sentiment analysis based on deep learning: A comparative study," *Electron.*, vol. 9, no. 3, p. 483, Mar. 2020, doi: [10.3390/electronics9030483](https://doi.org/10.3390/electronics9030483).
- [19] P. Singhal and P. Bhattacharyya. (2016). *Sentiment Analysis and Deep Learning a Survey*. [Online]. Available: <https://api.semanticscholar.org/CorpusID:38100008>
- [20] A. B. Sayeed, J. Boyd-Graber, B. Rusk, and A. Weinberg, "Grammatical structures for word-level sentiment detection," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Human Lang. Technol. (NAACL HLT)*. Stroudsburg, PA, USA: Association for Computational Linguistics, 2012, pp. 667–676.
- [21] C. Nicholls and F. Song, "Improving sentiment analysis with part-of-speech weighting," in *Proc. Int. Conf. Mach. Learn. Cybern.*, vol. 3, Jul. 2009, pp. 1592–1597.
- [22] K. Meichanetzidis, A. Toumi, G. de Felice, and B. Coecke, "Grammar-aware sentence classification on quantum computers," 2020, *arXiv:2012.03756*.
- [23] K. Meichanetzidis, A. Toumi, G. de Felice, and B. Coecke, "Grammar-aware sentence classification on quantum computers," *Quantum Mach. Intell.*, vol. 5, no. 1, p. 10, Feb. 2023, doi: [10.1007/s42484-023-00097-1](https://doi.org/10.1007/s42484-023-00097-1).
- [24] R. Lorenz, A. Pearson, K. Meichanetzidis, D. Kartsaklis, and B. Coecke, "QNLP in practice: Running compositional models of meaning on a quantum computer," *J. Artif. Intell. Res.*, vol. 76, pp. 1305–1342, Apr. 2023, doi: [10.1613/jair.1.14329](https://doi.org/10.1613/jair.1.14329).
- [25] D. E. Deutsch, "Quantum computational networks," *Proc. Roy. Soc. London A, Math. Phys. Sci.*, vol. 425, pp. 73–90, Sep. 1989.
- [26] C. P. Williams and S. H. Clearwater, *Explorations in Quantum Computing*. London, U.K.: Springer, 1998.
- [27] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter, "Elementary gates for quantum computation," *Phys. Rev. A, Gen. Phys.*, vol. 52, no. 5, pp. 3457–3467, Nov. 1995, doi: [10.1103/physreva.52.3457](https://doi.org/10.1103/physreva.52.3457).
- [28] N. A. Gershenfeld, *The Nature of Mathematical Modeling*. Cambridge, U.K.: Cambridge Univ. Press, 1999.
- [29] S. Sim, J. Romero, J. F. Gonthier, and A. A. Kunita, "Adaptive pruning-based optimization of parameterized quantum circuits," *Quantum Sci. Technol.*, vol. 6, no. 2, Apr. 2021, Art. no. 025019.
- [30] D. Shepherd and M. J. Bremner, "Temporally unstructured quantum computation," *Proc. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 465, no. 2105, pp. 1413–1439, Feb. 2009, doi: [10.1098/rspa.2008.0443](https://doi.org/10.1098/rspa.2008.0443).

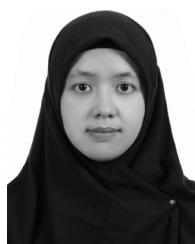


MUHAMMAD RIFAT ABIWARDANI received the B.Sc. degree (cum laude) in informatics from Institut Teknologi Bandung, Indonesia, in 2023. His current research interests include quantum machine learning and quantum NLP.



INFALL SYAFALNI (Member, IEEE) received the B.Eng. degree in electrical engineering from Institut Teknologi Bandung (ITB), Bandung, Indonesia, in 2008, the M.Sc. degree in electronics engineering from the University of Science Malaysia (USM), Penang, Malaysia, in 2011, and the Dr.Eng. degree in engineering from the Kyushu Institute of Technology (KIT), Iizuka, Japan, in 2014. From 2014 to 2015, he held a research position with KIT. From 2015 to 2018,

he was an ASIC Engineer with the ASIC Development Group, Logic Research Company Ltd., Fukuoka, Japan. In 2019, he joined ITB, where he is currently an Assistant Professor with the School of Electrical Engineering and Informatics and a Researcher with the University Center of Excellence on Microelectronics. In 2023, he is also currently a Visiting Scholar with the System Technology Co-Optimization (STCO) Program, Interuniversity Microelectronics Centre (IMEC), Leuven, Belgium. His research interests include logic synthesis, logic design, VLSI design, and efficient circuits and algorithms.



HARASHITA TATIMMA LARASATI (Graduate Student Member, IEEE) received the B.S. and M.S. degrees in telecommunication engineering from Institut Teknologi Bandung (ITB), Bandung, Indonesia, in 2016 and 2017, respectively. She is currently pursuing the Ph.D. degree in computer engineering with Pusan National University, Busan, Republic of Korea. She is also a Junior Lecturer with ITB. Her research interests include quantum computing and cryptanalysis, quantum machine learning, AI security, and networking.

machine learning, AI security, and networking.



RAHMAT MULYAWAN (Member, IEEE) received the B.Eng. degree in EE from ITB, Indonesia, in 2008, and the M.Sc. degree in EE from TU Delft, The Netherlands, in 2011. He is a member of the Microelectronics Centre, ITB. His research interests include intelligent signal processing, MIMO systems, and transceiver design for optical wireless communications.

...



FARISKA ZAKHRALATIVA RUSKANDA (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in natural language processing from the School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Bandung, Indonesia. She is currently an Assistant Professor with the Informatics Research Group, Institut Teknologi Bandung.